Table of Contents

Synthesis and purification of 4a-d.

The methyl esters **4a-d** were prepared as previously reported,[1] with acetyl-d_3-chloride (99 atom % D, Sigma) used for the synthesis of the deuterated substrate **4d**. **4a-d** were purified by preparative HPLC on a Phenomenex Luna C18(2) 100A AX1A Pa column (250.0 x 21.2 mm, 10 μm) using an Agilent HP 1200 HPLC. Mixtures of water and methanol (HPLC grade, both with 0.1% formic acid) were used with a flow rate of 20.0 mL/min and an elution gradient starting from 100 % water and linearly increasing to 100 % acetonitrile over 30 min; with UV detection set at 254, 280 and 210 nm. **4a-b** eluted at 11.0 min, whereas **4c-d** eluted at 13.6 min. The NMR data for **4a-c** corresponded to those previously reported.[1] For **methyl 6-(acetamido-d_3)-3-oxohexanoate** (**4d**): δ_H (400MHz, $CDCl_3$, Me_4Si) 1.81 (2H, quintet, *J* 7.0, CH_2), 2.61 (2H, t, *J* 6.5, CH_2), 3.25 (2H, apt q, *J* 6.5, *J* 7.0, CH_2), 3.47 (2H, s, CH_2), 3.75 (3H, s, OMe), 6.00 (1H, br s, NH); δ_C (100 MHz, $CDCl_3$ Me_4Si) 23.0 (t), 23.1 (q, *C*D_3), 38.8 (t), 40.3 (t), 48.9 (t), 52.4 (q, O*C*H_3), 167.8 (s), 170.9 (s), 202.7 (s, CH_2*C*OCH_2); *m*/*z* (EI) 205.1258 ($[M+H]^+$. $C_9H_{13}D_3NO_4$ requires 205.1262).

Construction of ΔDEBS 1, ΔDEBS 2 and ΔDEBS 3 mutant strains of *S. erythraea.*

Strains YD7 (ΔDEBS 1), YD69 (ΔDEBS 2) and YD13 (ΔDEBS 3) were obtained by in-frame deletions of *ery AI* (DEBS 1), *eryAII* (DEBS 2) and *eryAIII* (DEBS 3) respectively in *Saccharopolyspora erythraea* BIOT-1174 overproducing strain.[2]

The in-frame deletions were performed by using REDIRECT[3] which is based on the Red/ET recombineering technology,[4] with the following modifications.

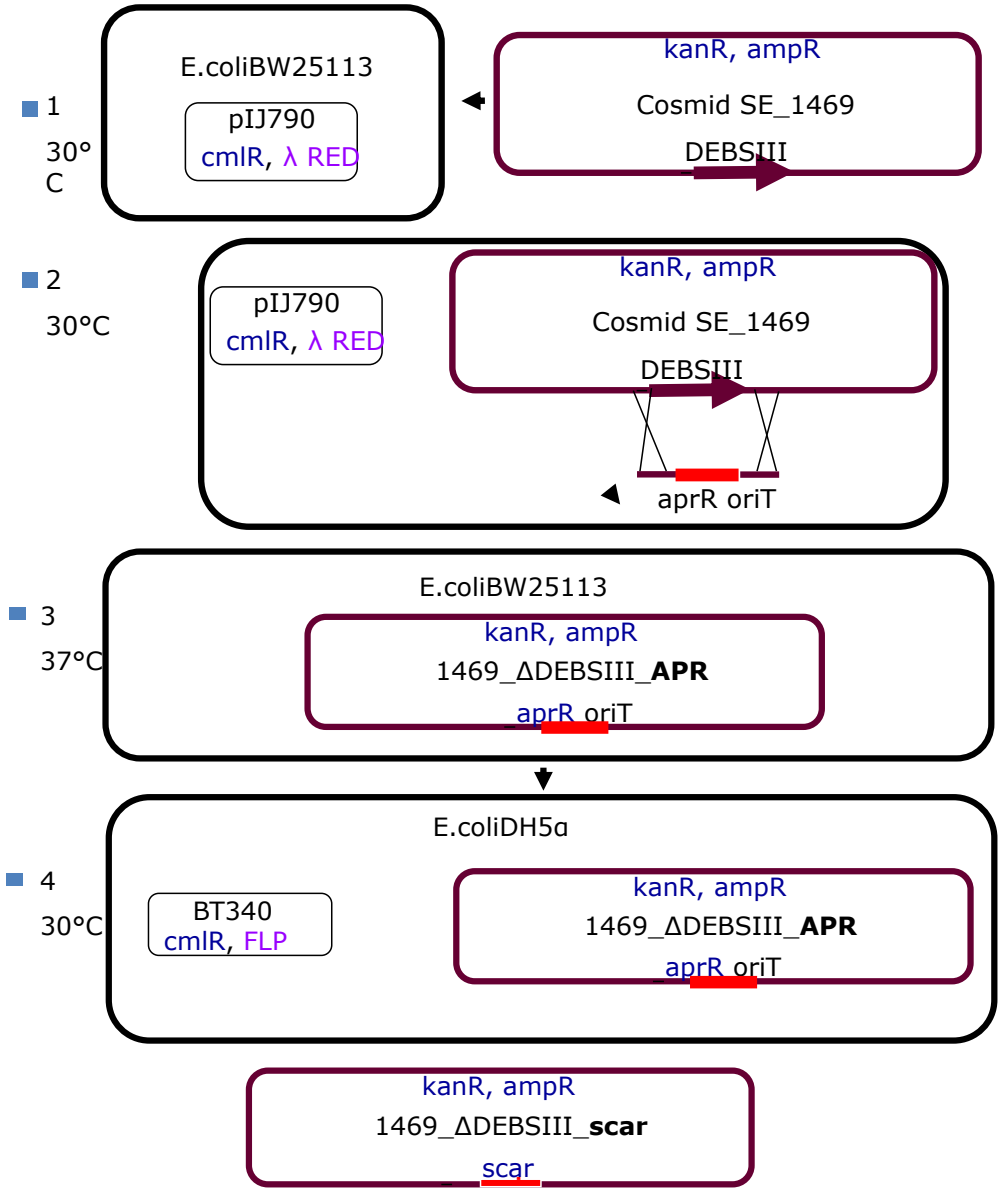
Suitable cosmids containing each targeted gene were identified from a genomic library of *S. erythraea* BIOT-1714. These cosmids harbouring respectively *eryAI* (DEBS 1) (SE_P1611); *eryAII* (DEBS 2) (SE_P533); or *eryAIII* (DEBS 3) genes (SE_P1469) were used to transform *E. coli* strain BW25113 by electroporation. BW25113 harbours plasmid pIJ790 which encodes λ RED recombinase and a chloramphenicol resistance marker. Transformants were selected by their resistance to both kanamycin (selection for the cosmid) and chloramphenicol. The PCR fragment – or disruption cassette - harbouring the *aprR* gene and the origin of plasmid transfer *oriT*, and flanked by 60bp fragments with homology to

[1] M. Tosin, L. Betancor, E. Stephens, W. M. A. Li, J. B. Spencer, P. F. Leadlay, *ChemBioChem* 2010, **11**, 539.
[2] Kind gift of Biotica Technology Ltd.
[3] B. Gust ,G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc Natl Acad Sci USA* 2003, **100**, 1541.
[4] Y. Zhang, F. Buchholz, J. P. P. Muyrers, A.F. Stewart, *Nature Genetics*, 1998, **20**, 123.

sequences adjacent to the gene to be deleted, was produced according to the manufacturer's instructions. The disruption cassette was introduced by electroporation into strain pIJ790/ SE_P1611 and the resulting mutant strain P1611_ΔDEBSI_APR was selected by growth on kanamycin- and apramycin-containing plates. In this strain the *eryAI* (DEBS 1) gene was substituted by the *aprR*/*oriT* disruption cassette. The disruption cassette is flanked by FRT sites (recognition targets for Flp recombinase). *E. coli* DH5α cells containing the temperature-sensitive plasmid BT340 bearing the gene for FLP recombinase were transformed with cosmid DNA isolated from the P1611_ΔDEBSI_APR strain. FLP synthesis and loss of the plasmid are induced at 42 C. Expression of the FLP-recombinase in *E. coli* removes the central part of the disruption cassette, leaving a 81bp scar that lacks stop codons in preferred reading frame. This allowed generation of non-polar unmarked in-frame deletions of the desired gene (resulting in cosmid P1611_ΔDEBSI-scar). The *aprR*/*oriT* cassette was then introduced into the backbone of cosmid P1611_ΔDEBSI by Redirect technology[3] resulting in 1611_ΔDEBSI_scat (SuperCos Apr OriT) cosmid, which allowed introduction of the delivery vector into *S. erythraea* via conjugation from *E. coli* strain ET12567/pUZ 8002. Following conjugation, *S. erythraea* transformants containing the delivery cosmid 1611_ΔDEBSI_scat integrated into the genome (first crossover) were selected by growth on kanamycin- and apramycin-containing plates. Second crossovers with complete deletion of the gene of interest were selected by subsequent growth on plates without kanamycin- and apramycin-selection. Each deletion and its location were confirmed by PCR and by sequencing of the PCR fragment in the resulting strain YD7. In this strain *eryAI* was substituted in-frame by an 81bp scar sequence lacking start and stop codons. The same procedure was followed to make Δ eryAII and Δ eryAIII mutant strains (YD69 and YD13 respectively; refer to Figure 1S for YD13 construction).

Figure 1S (following two pages):

Construction of the ΔDEBS 3 mutant strain of *S. erythraea*



- FLP-recombinase removes the central part of disruption cassette, leaving 81bp **scar** that lacks stop codons

20bp +19bp priming sequence + 42bp FLP core recombination site

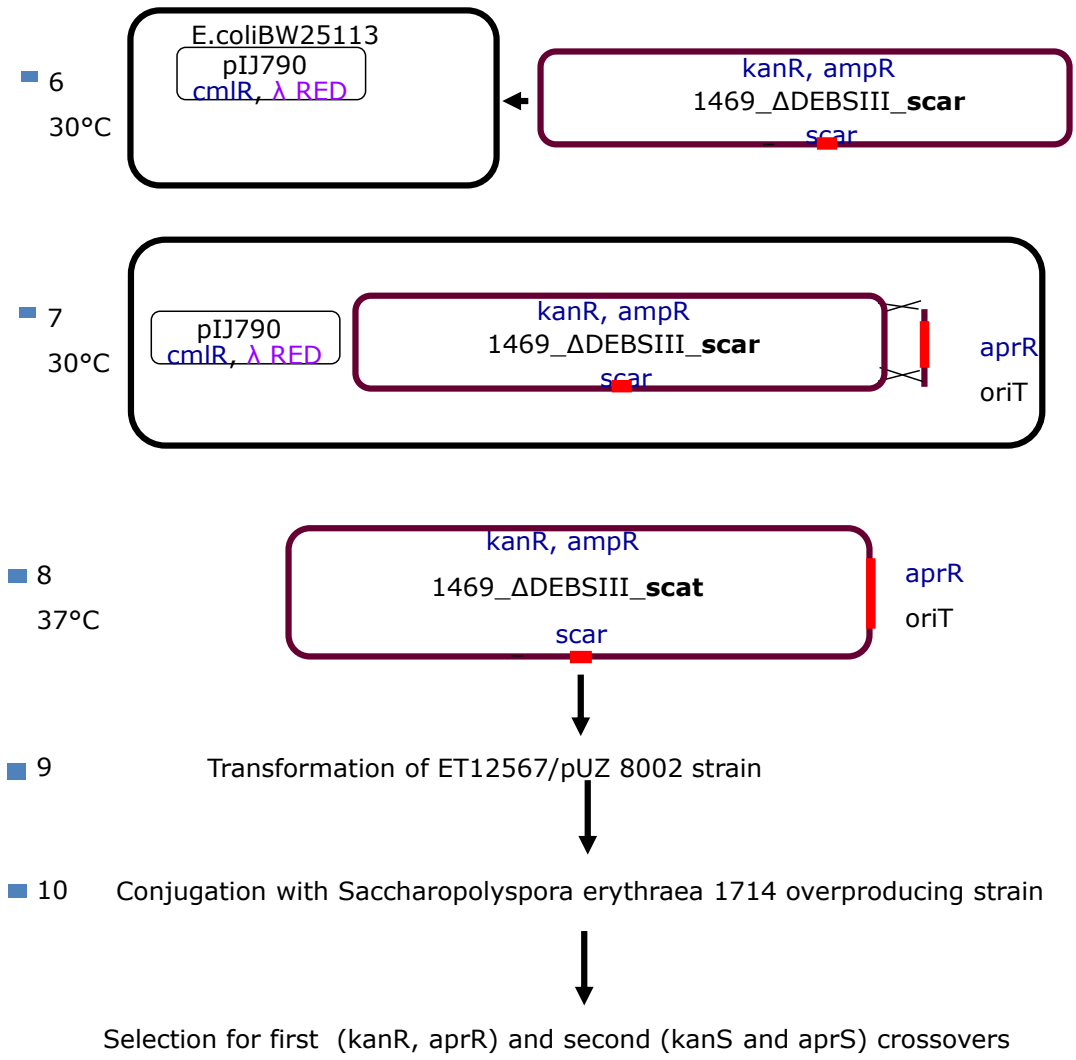
AG**ATG**ATTCCGGGGATCCGTCGACCTGCAGTTCGAAGTTCCTATTCTCTAGAAAGTA
TAGGAACTTCGAAGCAGCTCCAGCCTACA**TGA**CCAC

↓

- 5 — 42°C

Induce FLP synthesis and loss of temperature sensitive FLP recombination plasmid BT340;

Plasmid isolation, sequencing



Construction of DEBS 1-DEBS 2(DH⁻)-TE *S. erythraea*

S. erythraea DEBS 1-DEBS 2(DH⁻)-TE was obtained by transformation of *S. erythraea* DJB[5] with a bifunctional plasmid (pUC18-pIJ702) containing an insert from *ery AII* (the region between KR and ACP active sites) ligated to the end of *ery AIII* (from ACP2 to the end).
For further information email pfl10@mole.bio.cam.ac.uk.

[5] D. J. Bevitt, PhD thesis, University of Cambridge, **1992**.

Growth of *S. erythraea* strains and mass spectrometry analysis.

All the strains, with the exception of DEBS 1-DEBS 2(DH$^-$)-TE *S. erythraea* (for which tap-water medium, TWM,[6] was used throughout), were grown in M79 medium (10 mL) for 3 days at 30ºC. The seed cultures (100 μl aliquots) were plated onto TWM-agar (5 mL, in duplicate/triplicate copy) containing **4a-d** in 10 mM concentration. Control plates in absence of **4a-d** or the strains were prepared, as well as YD7 (ΔDEBS 1) plates as negative controls.
After incubation at 30 ºC for 4 days the agar plates were extracted twice with ethyl acetate (10 mL x 2). The extracts were concentrated and the residues were redissolved in HPLC-grade methanol (1 mL) for mass spectrometry analysis.
HPLC-HR-ESI-MS analyses of *S. erythraea* extracts were performed on a Thermo Electron LTQ-Orbitrap. Samples (1- 10 μl) were injected onto a Dionex Acclaim C18 PepMap 100 column (150 mm x 1.0 mm, 3 μm), eluting with a linear gradient of 0 % to 100 % B in 28 min with a flow rate of 50 μl/min (A: 98 % H_2O, 2 % MeCN, 0.1 % formic acid, B: 90 % MeCN, 10 % H_2O, 0.1 % formic acid). The mass spectrometer was run in positive ionization mode, scanning from *m/z* 100 to 1800, with the FTMS analyser resolution set at 60K. Selected ion search within 5 ppm was performed, as well as high resolution fragmentation (collision energy set to 30-35%) for the putative biosynthetic intermediates.

[6] C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton, P. F. Leadlay, *Gene*, 1998, **216**, 215.

In vivo hydrolysis of 4c

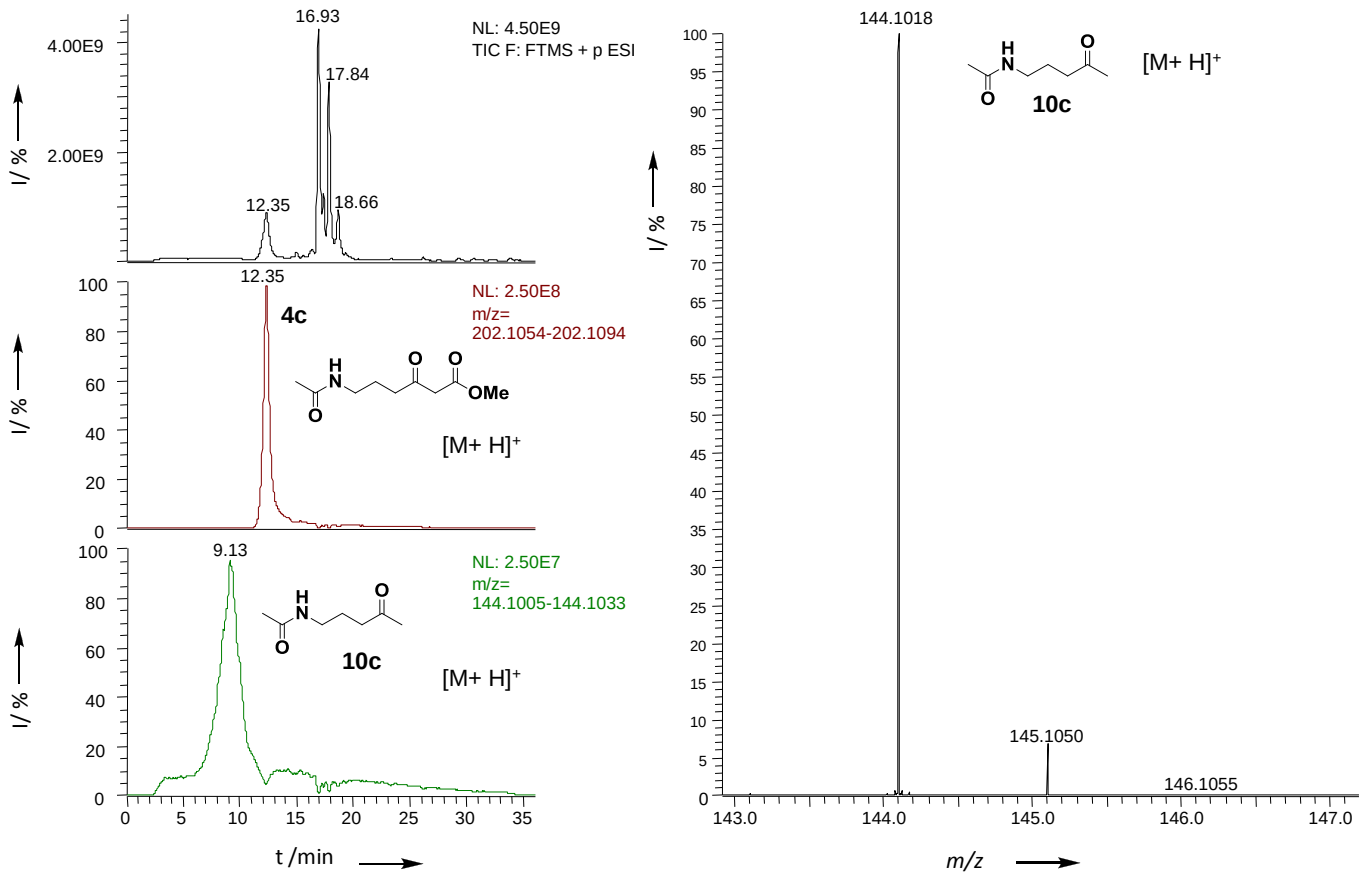
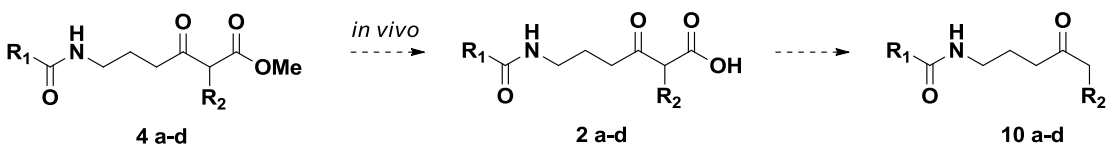


Figure 2S: LC-ESI-HRMS analysis of the organic extract of *S. erythraea* BIOT-1714 grown in presence of methyl-6-acetamido-3-oxohexanoate **4c** (10 mM). The Total Ion Current and the $[M+H]^+$ extracted ion traces (5 ppm mass accuracy) for the ester **4c** and the *N*-(4-oxopentyl)acetamide **10c** (*m/z* 144.1018 found, $C_7H_{14}NO_2^+$), deriving from ester hydrolysis to **2c** and decarboxylation (below), are shown.



In the TIC trace the peaks at R_T = 16.9, 17.8 and 18.7 min correspond to erythromycin A and other erythromycin derivatives respectively.

Off-loading of diketides from Δ DEBS 2 *S. erythraea via* 4a and 4b

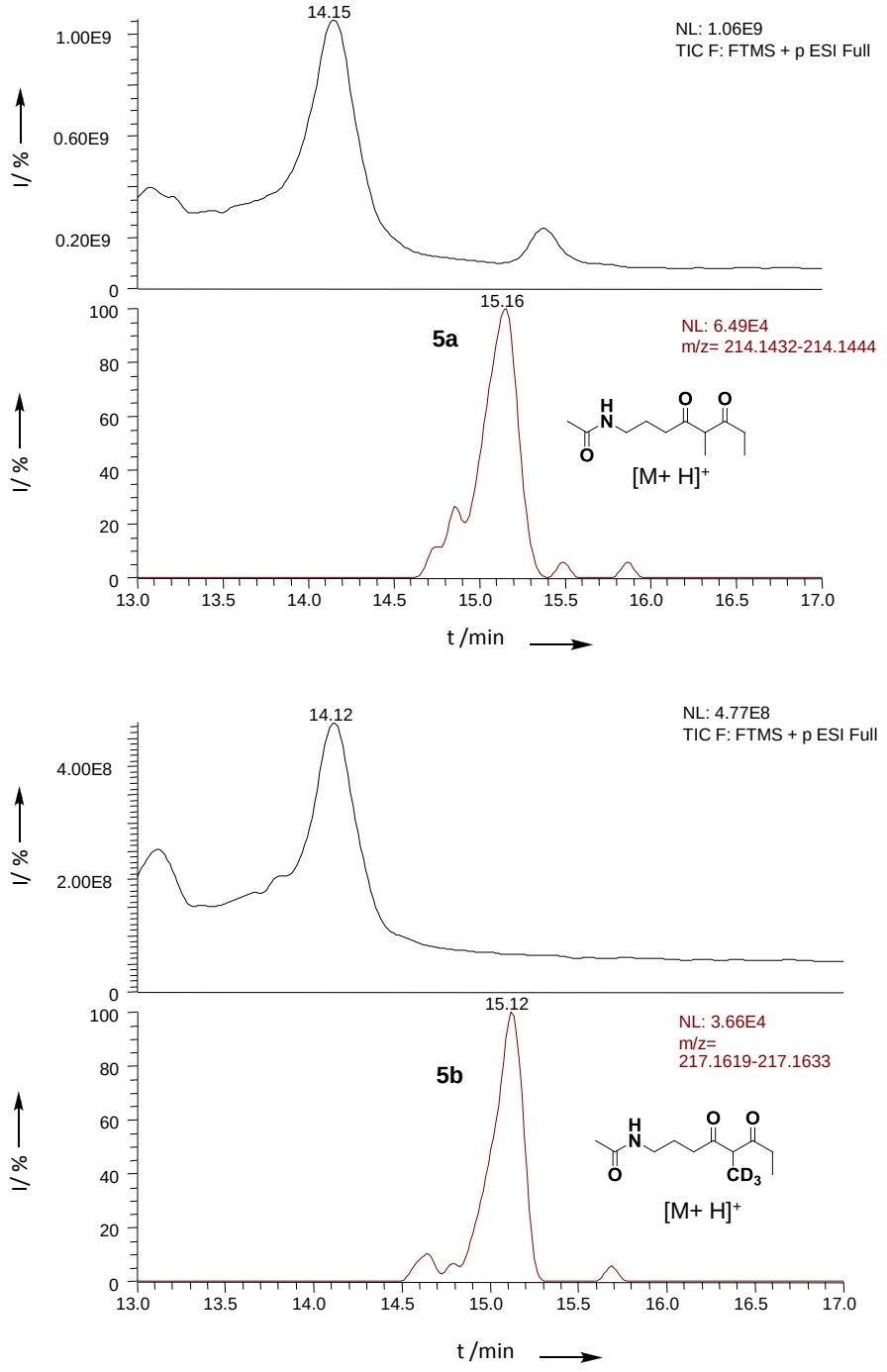


Figure 3S: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* Δ DEBS 2 grown in presence of methyl-6-acetamido-2-methyl-3-oxohexanoate **4a** (10 mM, top trace) and its 2-(methyl-d_3) analogue **4b** (10 mM, bottom trace). In the TIC traces the peaks at R_T = 14.1 min correspond to **4a** and **4b** respectively. The Total Ion Current and the $[M+H]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded diketides **5a-b** are shown. The HR-MS^2 fragmentation for **5a** and **5b** is identical to that previously reported for analogous diketide species off-loaded from a DEBS module *in vitro*.[1]

Off-loading of diketides from Δ DEBS 3 *S. erythraea via* 4c and 4d

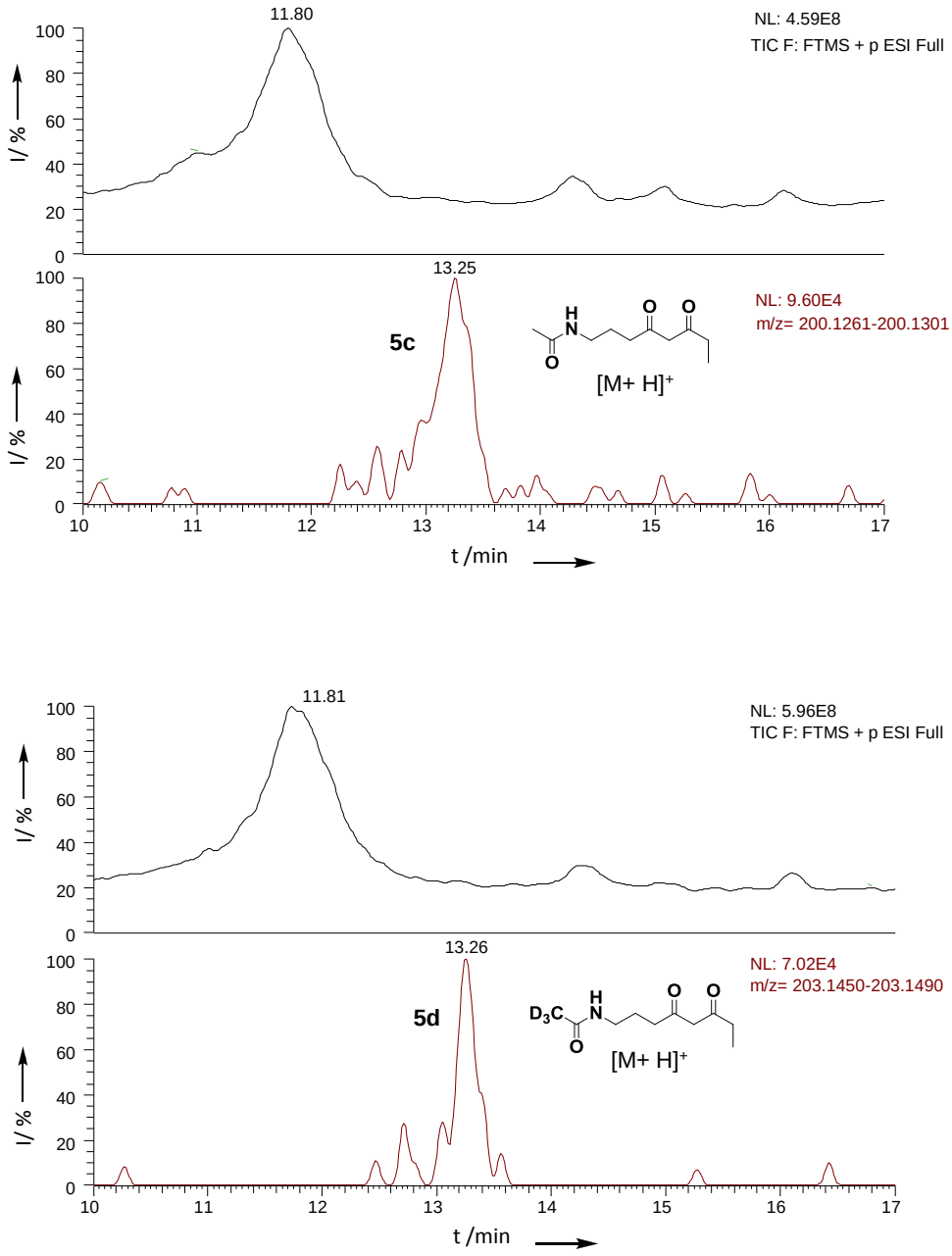


Figure 4S: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* Δ DEBS 3 grown in presence of methyl-6-acetamido-3-oxohexanoate **4c** (10 mM, top trace) and its acetamido-d_3 **4d** analogue (10 mM, bottom trace). In the TIC traces the peaks at R_T = 11.8 min correspond to **4c** and **4d** respectively. The Total Ion Current and the $[M+H]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded diketides **5c-d** are shown. The HR-MS^2 fragmentation for **5c** and **5d** is identical to that previously reported for analogous diketide species off-loaded from a DEBS module *in vitro*.[1]

Off-loading of triketides from Δ DEBS 2 *S. erythraea via* 4c and 4d

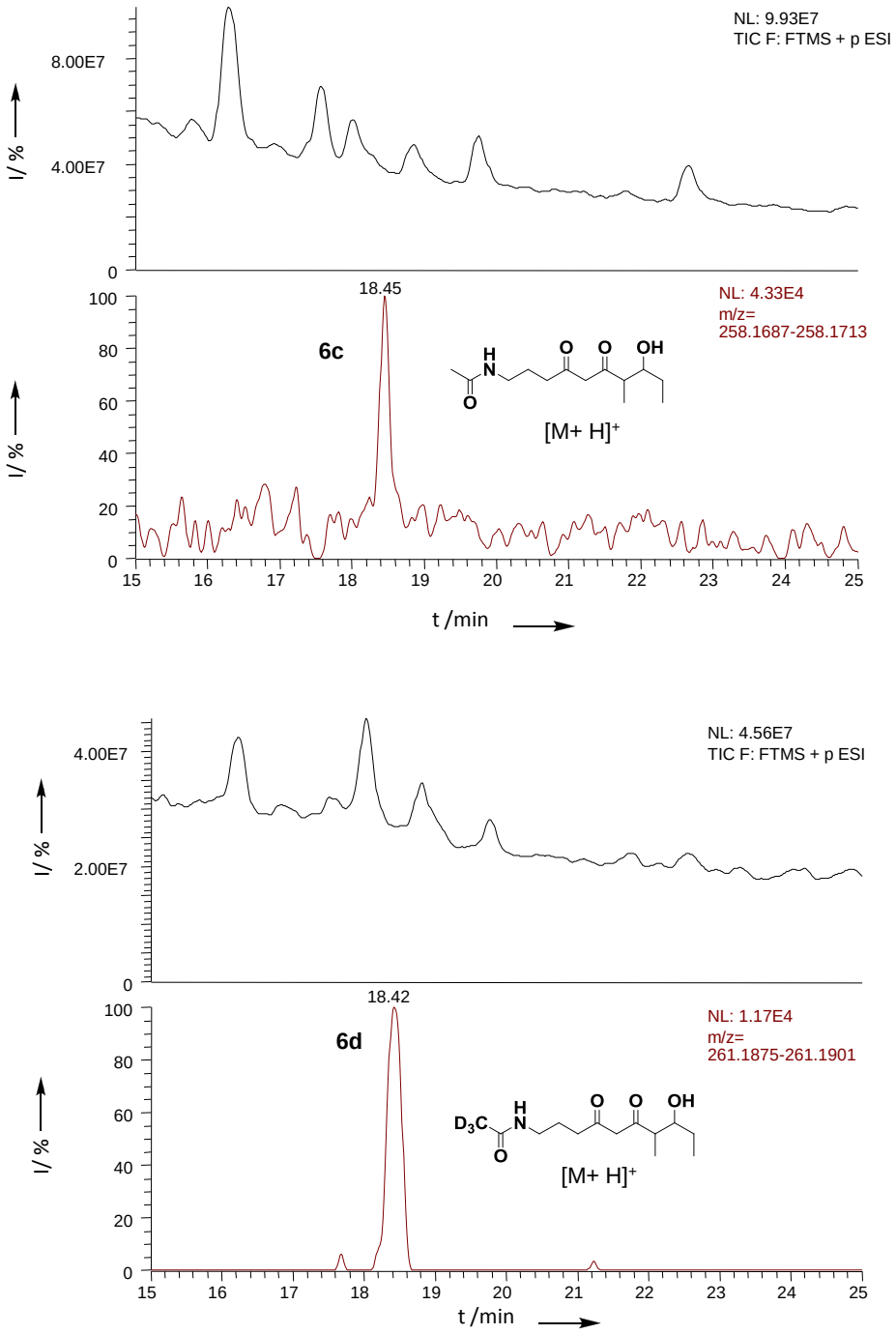


Figure 5S: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* Δ DEBS 2 grown in presence of methyl-6-acetamido-3-oxohexanoate **4c** (10 mM, top trace) and its acetamido-d_3 **4d** analogue (10 mM, bottom trace). The Total Ion Current and the $[M+H]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded triketides **6c-d** are shown.

HR-MS^2 characterization of triketides (6 c-d)

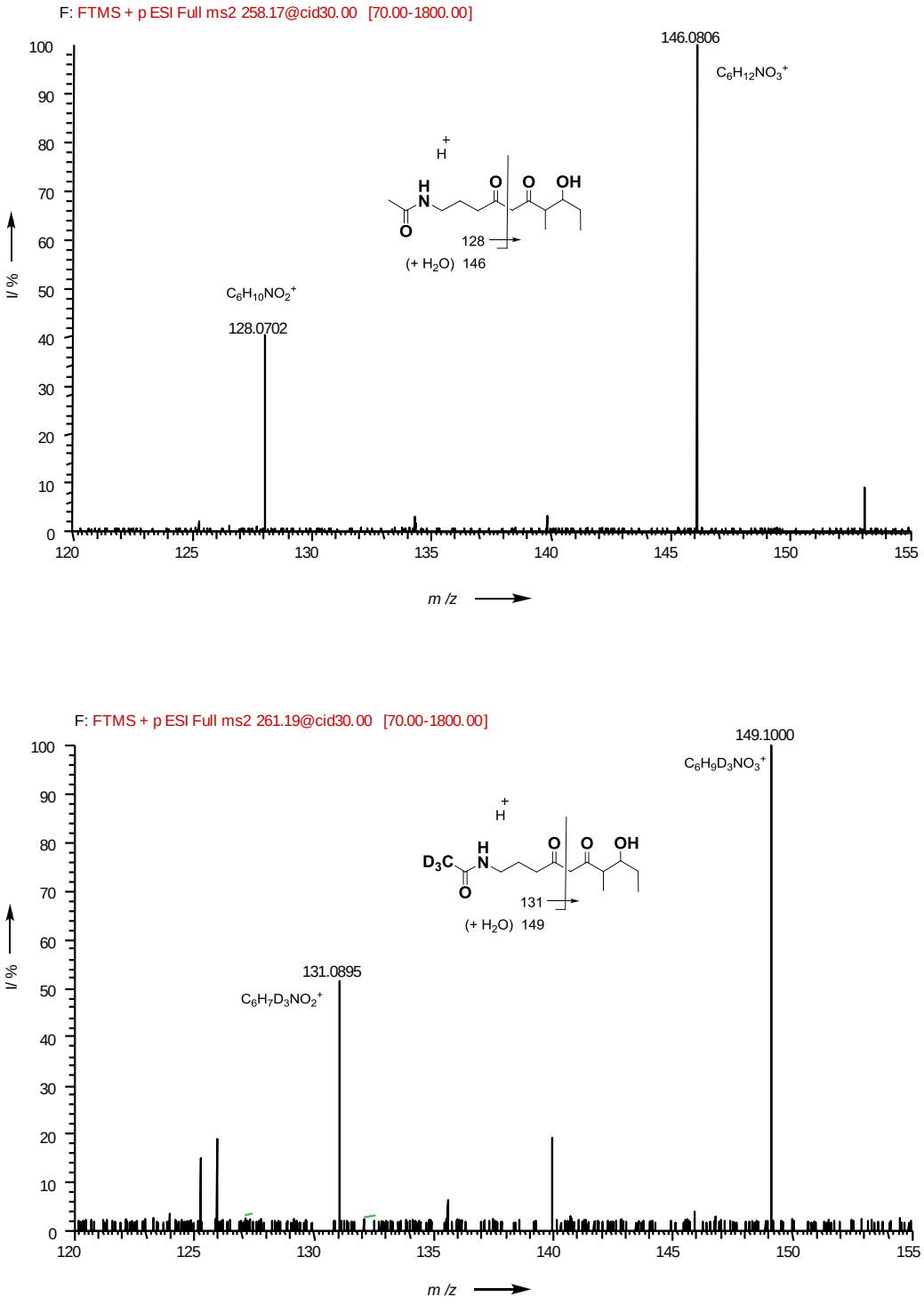


Figure 6S: HR-ESI-MS^2 analysis of the off-loaded triketides **6c** (top) and **6d** (bottom).

Off-loading and HR-MS^2 characterization of triketide 6a from Δ DEBS 2 *S. erythraea via* 4a

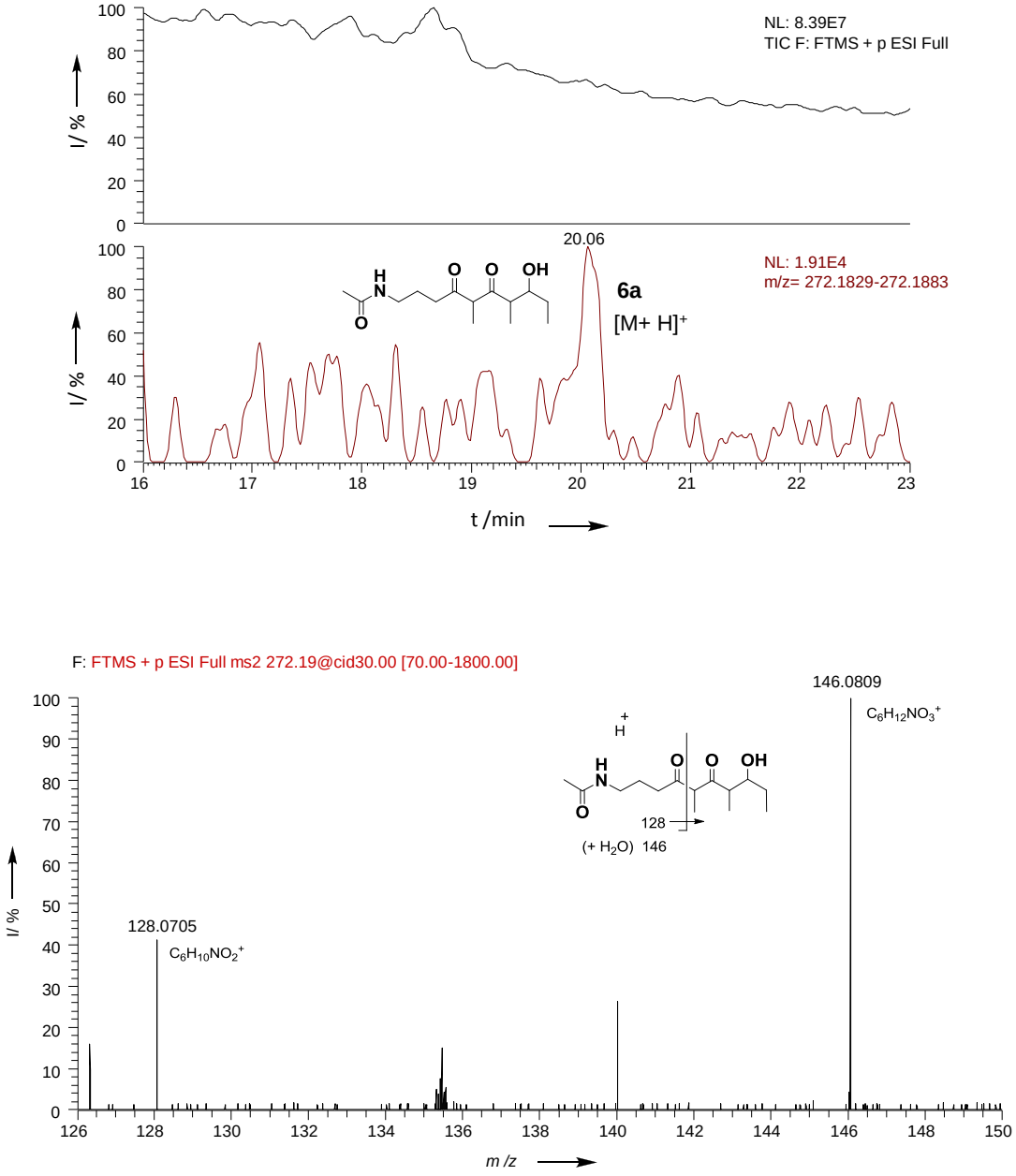


Figure 7S: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* Δ DEBS 2 grown in presence of methyl-6-acetamido-2-methyl-3-oxohexanoate **4a** (10 mM). The Total Ion Current and the $[M+H]^+$ extracted ion trace (5 ppm mass accuracy) for the off-loaded triketide **6a** are shown (top). Despite the low amount of **6a**, its HR-ESI-MS^2 analysis (below) is consistent with that of **6c** and **6d** (Figure 6S). Its deuterated counterpart **6b** has also been observed in minor amount in analogous experiments (Table 1, entry b).

Off-loading of tetraketides and pentaketides from Δ DEBS 3 *S. erythraea via* 4c and 4d

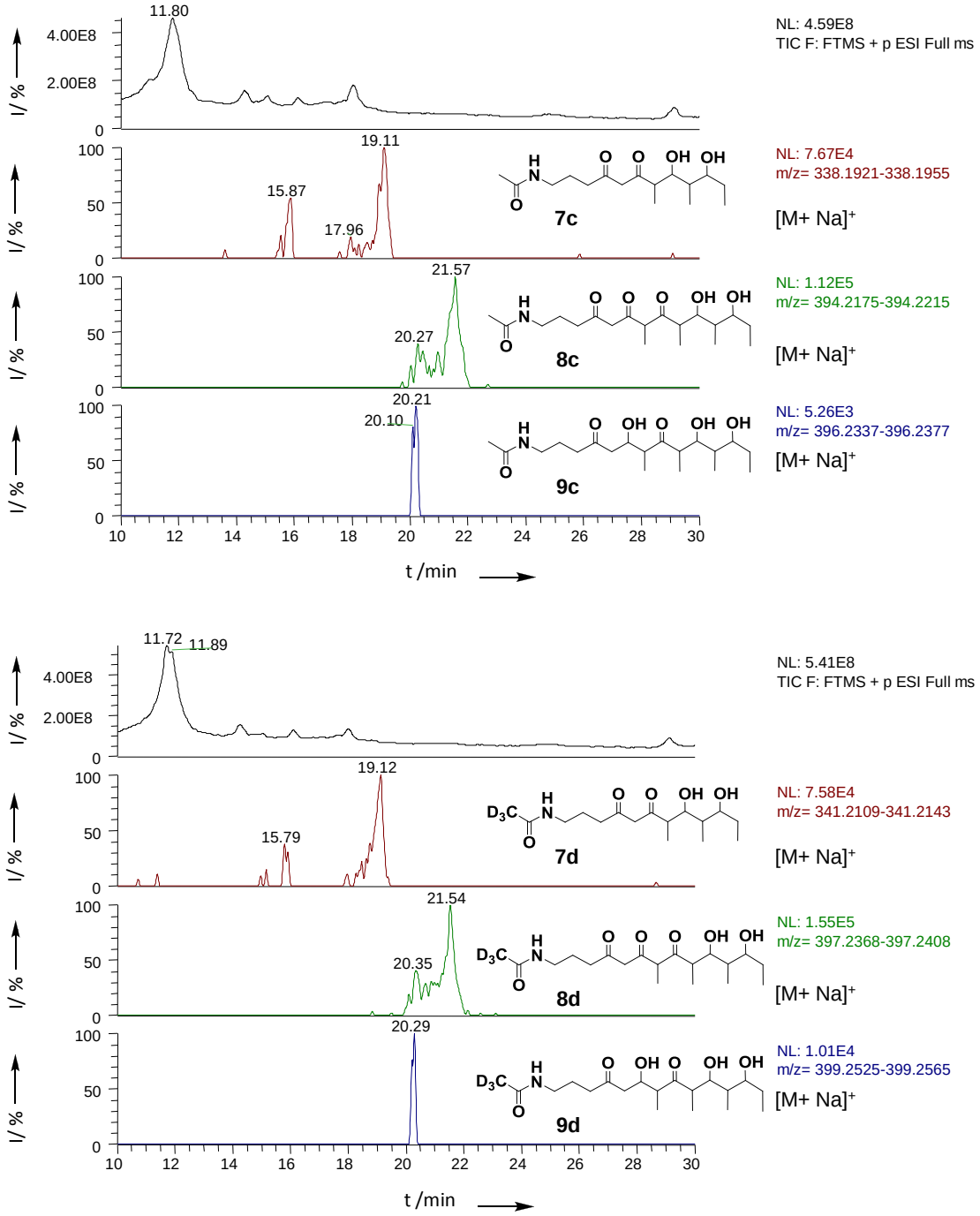


Figure 8S: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* Δ DEBS 3 grown in presence of methyl-6-acetamido-3-oxohexanoate **4c** (10 mM, top trace) and its acetamido-d_3 **4d** analogue (10 mM, bottom trace). The Total Ion Current and the $[M+Na]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded intermediates **7-9c-d** are shown. Two major distinct peaks with exact mass corresponding to the formula of **7c-d** and **8c-d** were detected.

Off-loading of tetraketides and pentaketides from DEBS 1- DEBS 2 (DH⁻)-TE *S. erythraea via* 4c and 4d

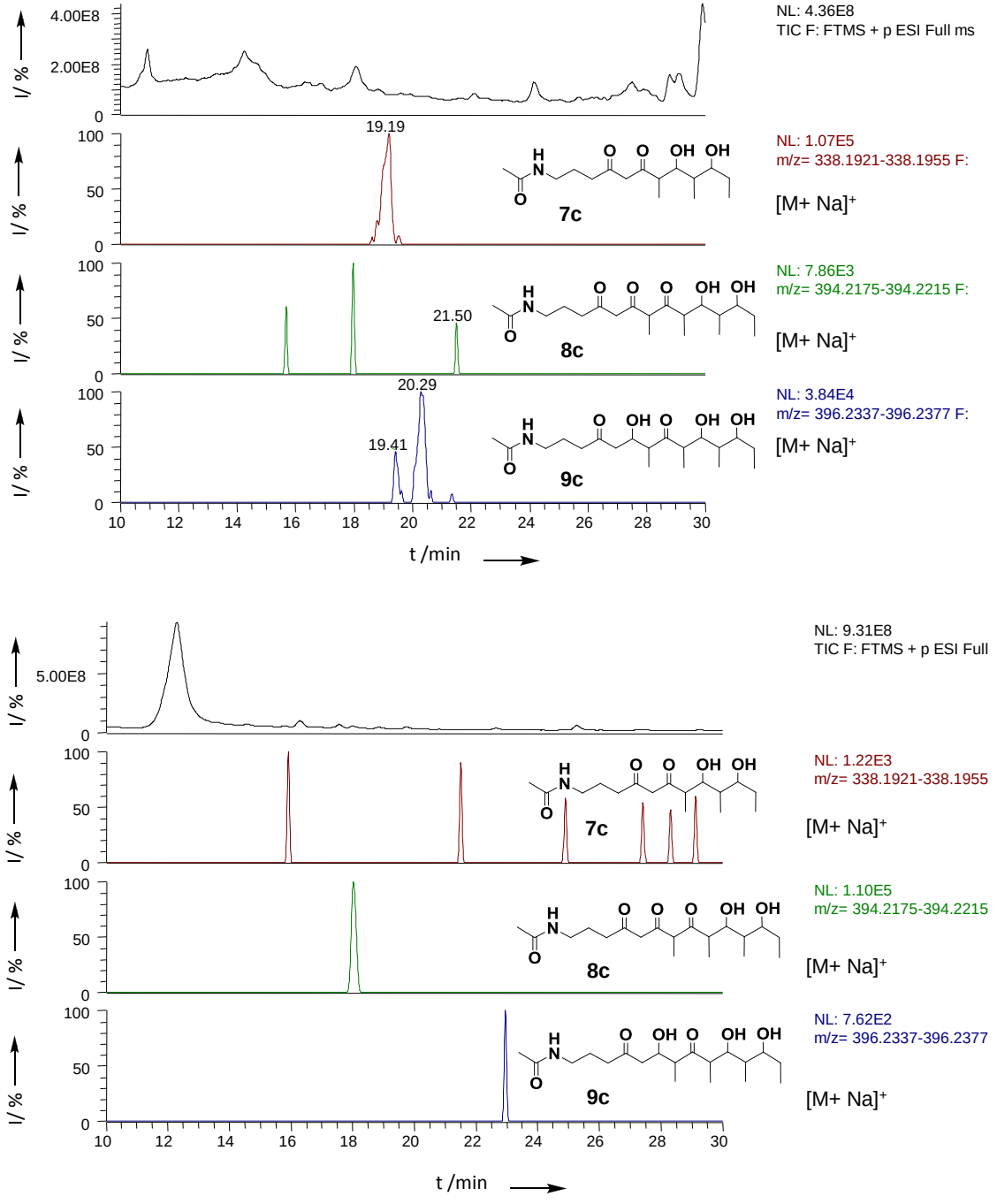


Figure 9S. Top: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* DEBS 1-DEBS 2 (DH^-)-TE grown in presence of **4c** (10 mM). The Total Ion Current and the $[M+Na]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded intermediates **7-9 c** are shown. Two major distinct peaks with exact mass corresponding to the formula of **9c** were detected. Bottom: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* ΔDEBS 2 grown as control in presence of **4c** (10 mM). No tetraketide or pentaketide intermediates were detected.

Off-loading of tetraketides and pentaketides from *S. erythraea* BIOT-1714

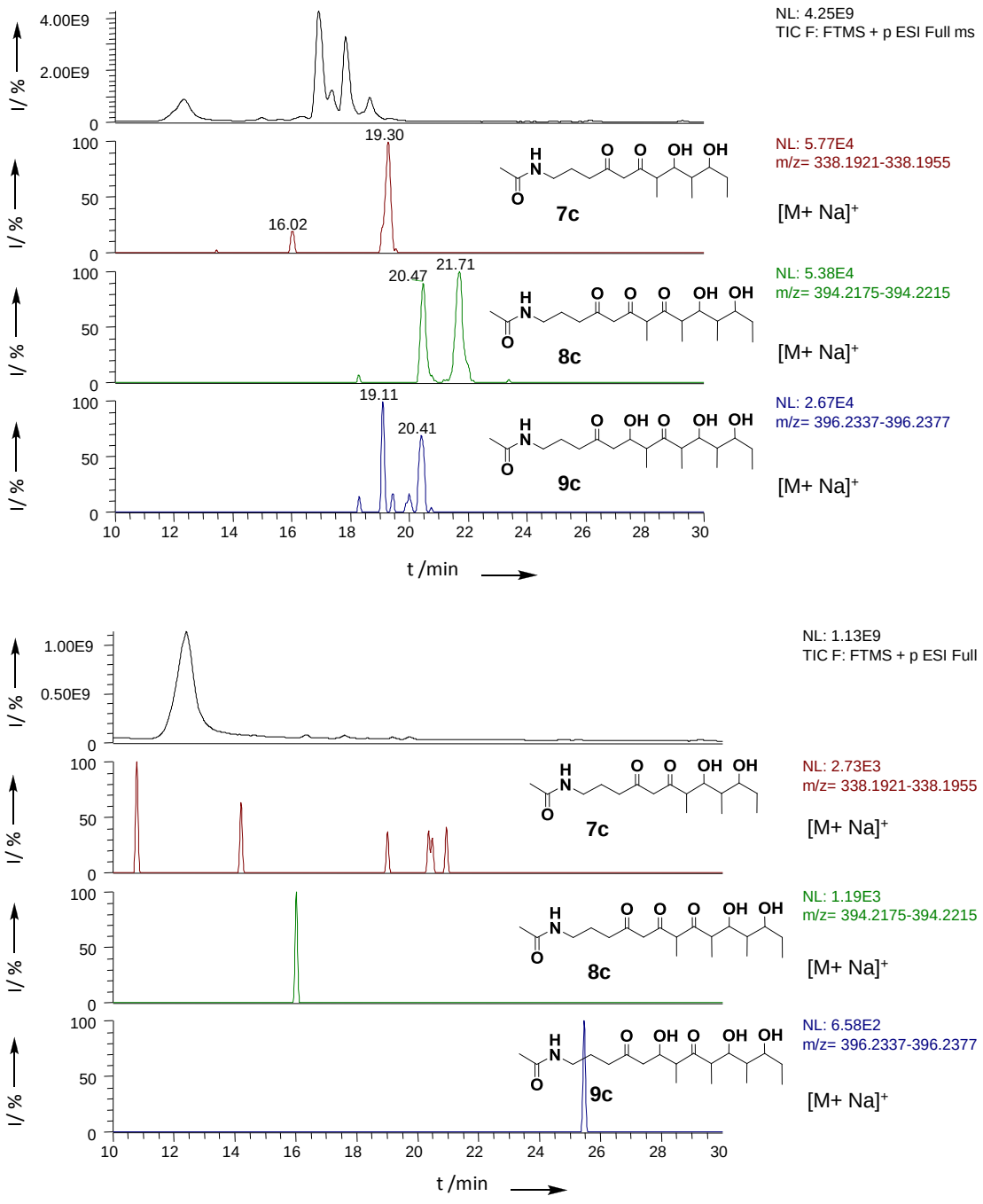


Figure 10S. Top: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* BIOT-1714 grown in presence of **4c** (10 mM). The Total Ion Current and the $[M+Na]^+$ extracted ion traces (5 ppm mass accuracy) for the off-loaded intermediates **7-9 c** are shown. Two major distinct peaks with exact mass corresponding to the formula of **8c** and **9c** were detected. Bottom: LC-ESI-HRMS analysis of the organic extracts of *S. erythraea* ΔDEBS 1 grown as control in presence of **4c** (10 mM). No tetraketide or pentaketide intermediates were detected.

HR-MS^2 characterization of tetraketides (7c-d)

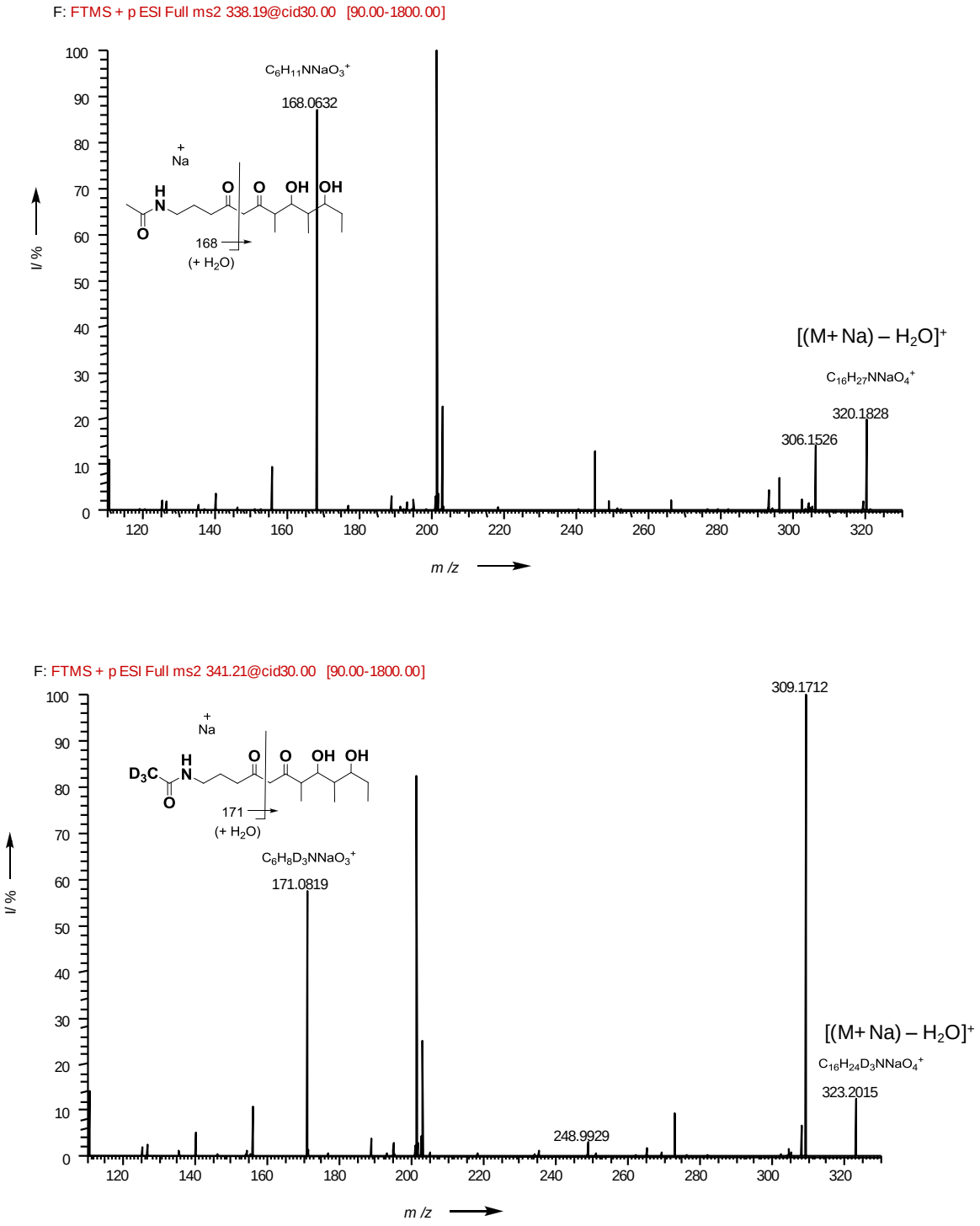


Figure 11S: HR-ESI-MS^2 analysis of the off-loaded tetraketides **7c** (top) and **7d** (bottom) at R_T = 19.1 min (refer to Fig 8S).

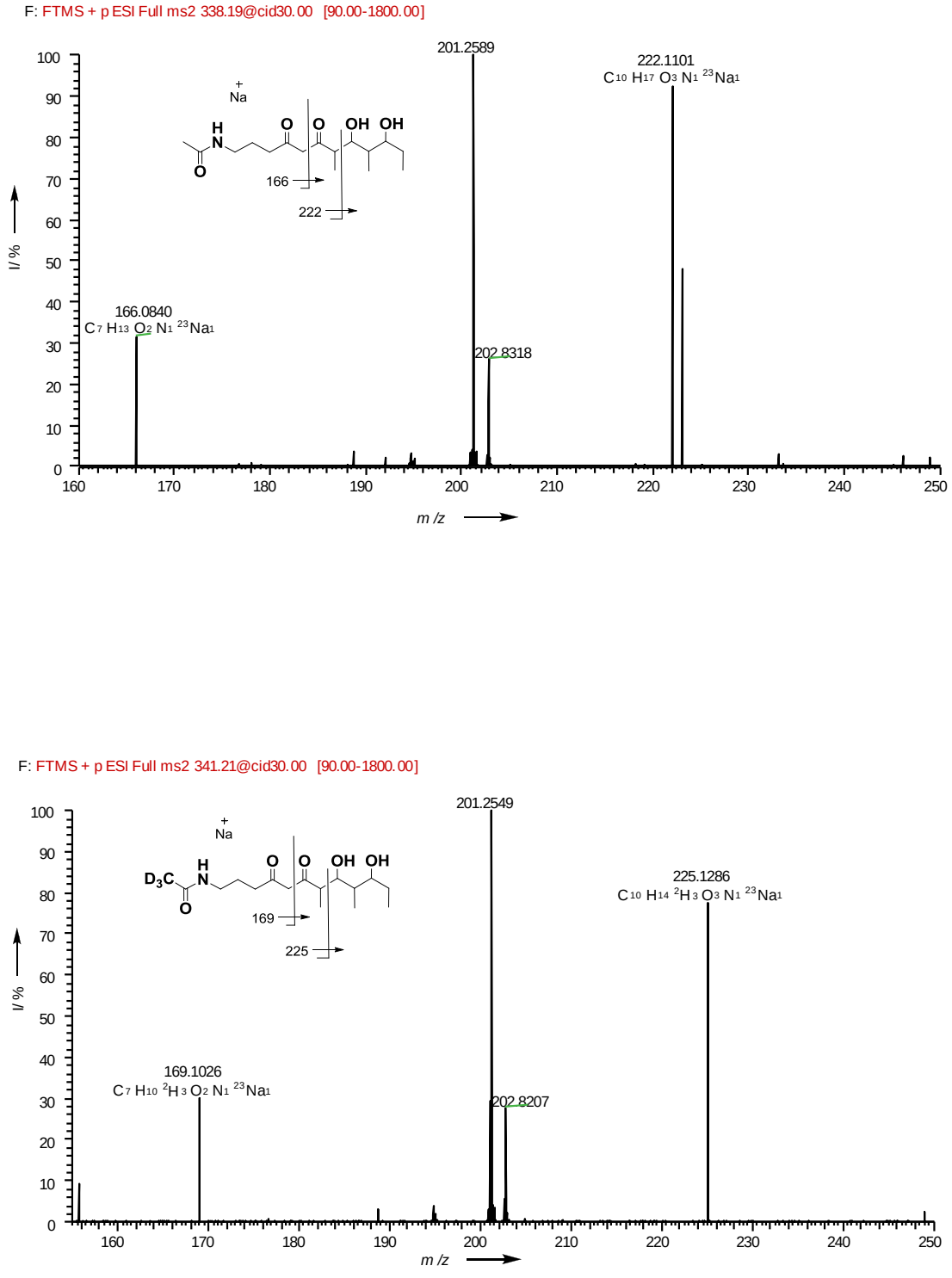
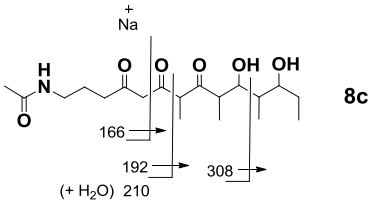
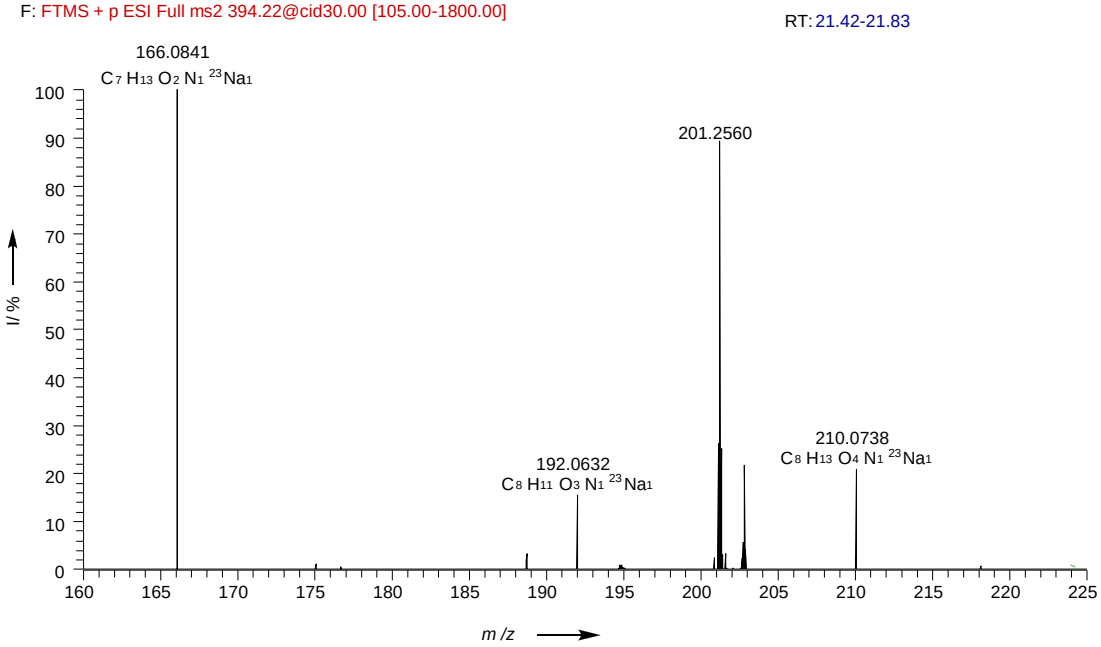


Figure 12S: HR-ESI-MS^2 analysis of the off-loaded tetraketides **7c** (top) and **7d** (bottom) at R_T = 15.8 min (refer to Fig 8S).

HR-MS^2 characterization of pentaketides (8c-d)



8c



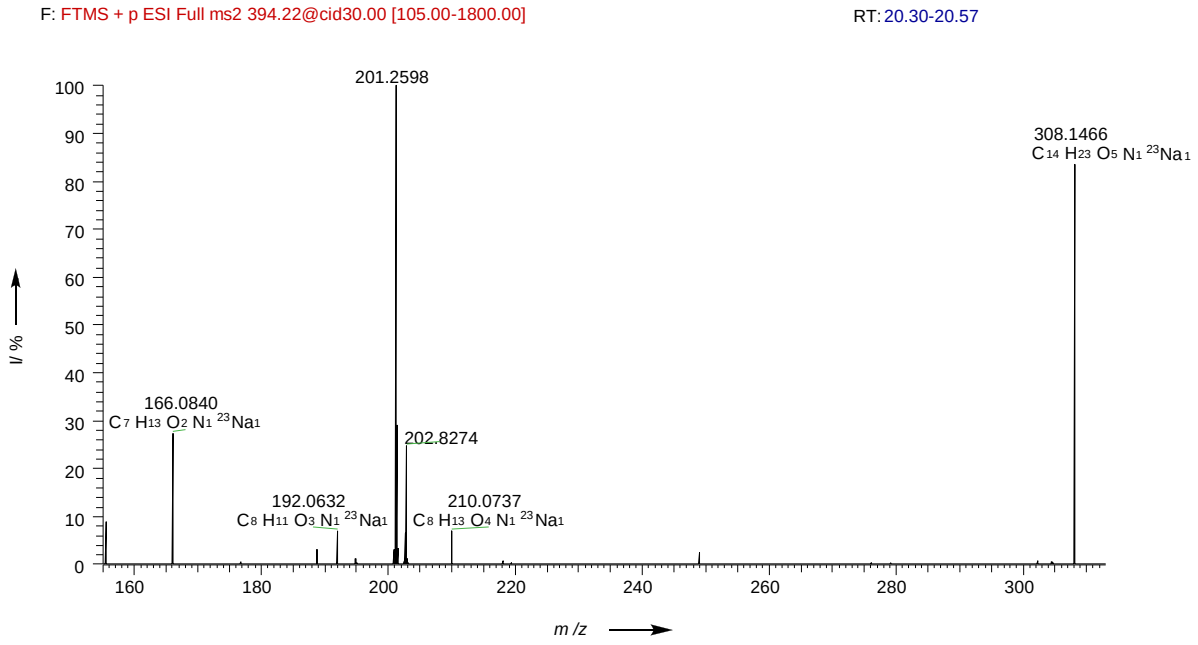
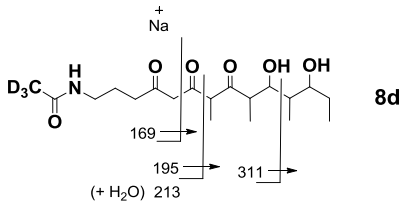
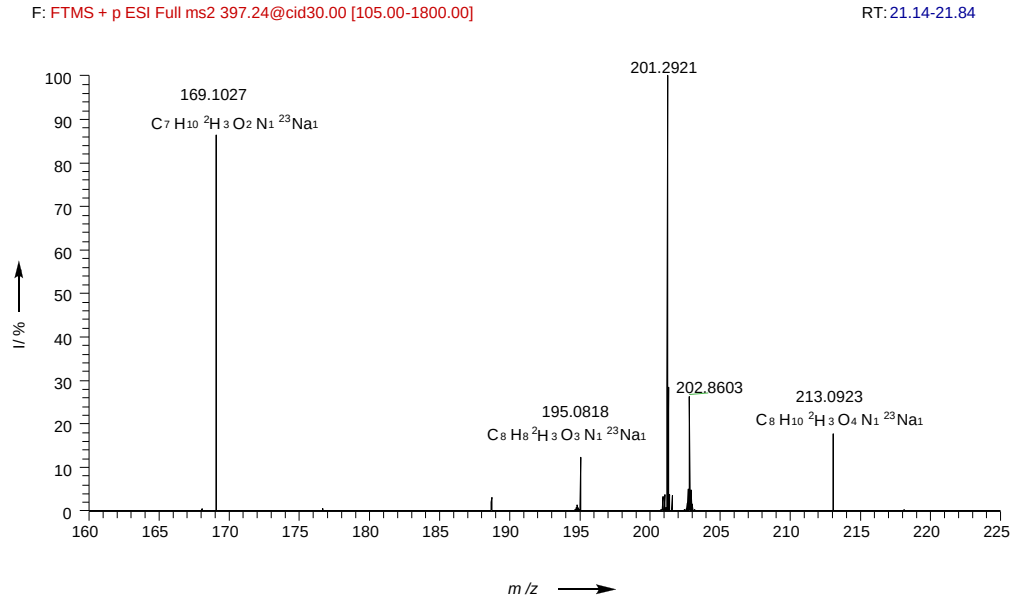


Figure 13S: HR-ESI-MS^2 analysis of the two distinct peaks identified for **8c** (R_T of 21.7 and 20.4 min, top and bottom, respectively; refer to Fig 8S and 10S). Multiple peaks could rise from partial cyclization and/or epimerization. Further work is in progress in order to establish their origin.





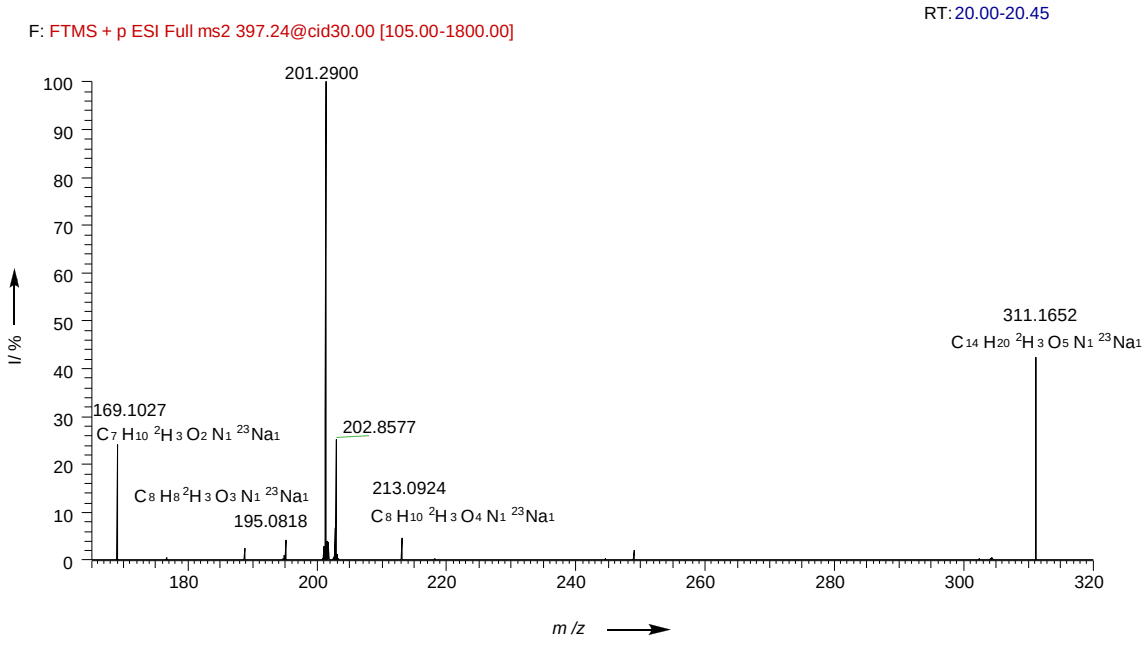


Figure 14S: HR-ESI-MS^2 analysis of the two distinct peaks identified for **8d** (R_T of 21.7 and 20.4 min, top and bottom, respectively, refer to Fig 8S). Multiple peaks could rise from partial cyclization and/or epimerization. Further work is in progress in order to establish their origin.

HR-MS^2 characterization of pentaketides (9c-d)

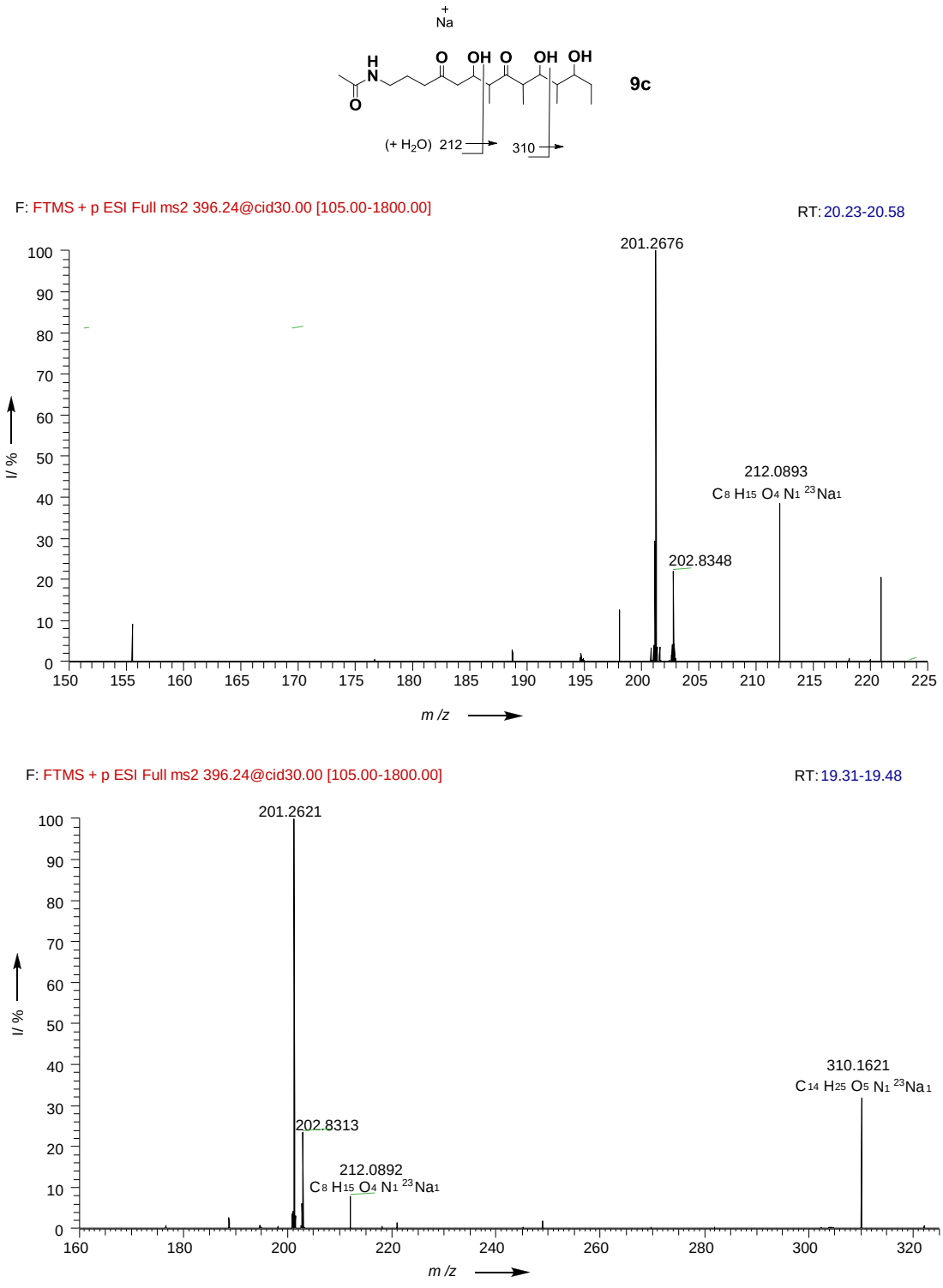
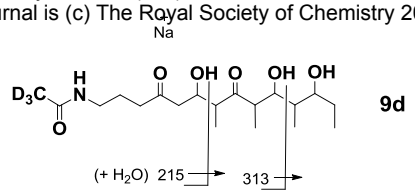
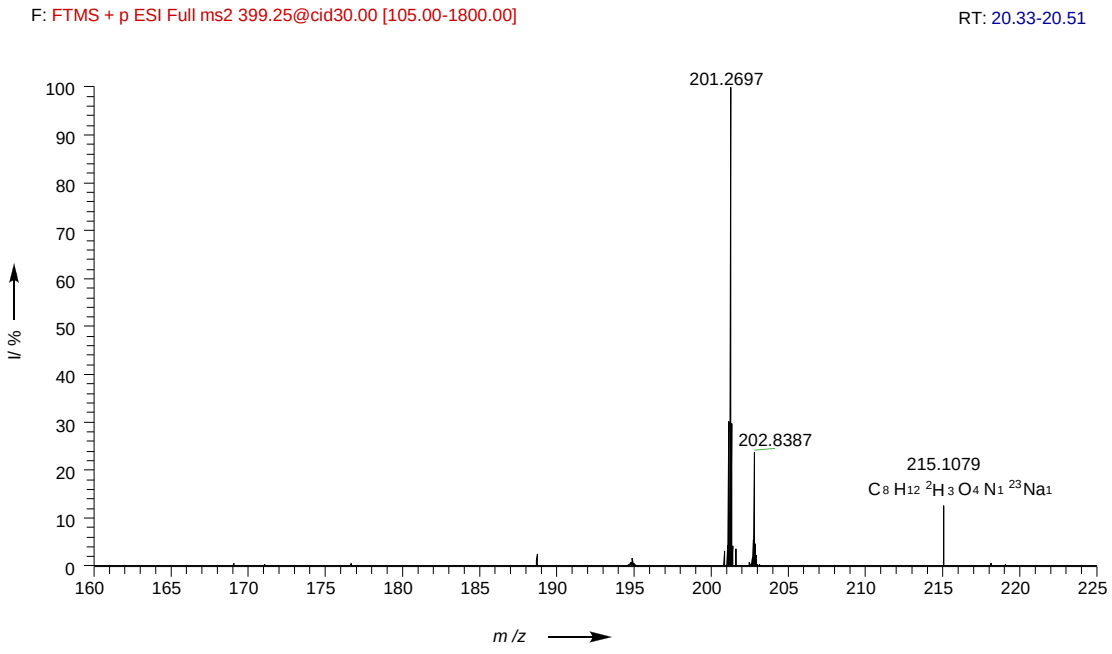


Figure 15S: HR-ESI-MS^2 analysis of the two distinct peaks identified for **9c** (R_T of 20.3 and 19.4 min, top and bottom, respectively; refer to Fig. 9S). Multiple peaks could rise from partial cyclization and/or epimerization. Further work is in progress in order to establish their origin.



9d



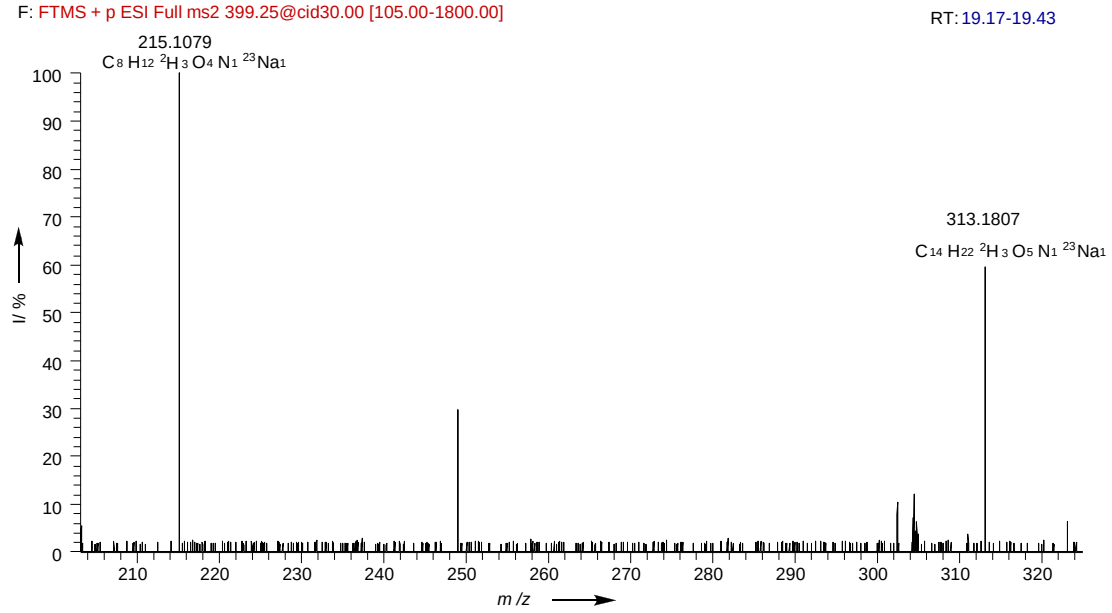


Figure 16S: HR-ESI-MS^2 analysis of the two distinct peaks identified for **9d** (R_T of 20.3 and 19.4 min, top and bottom, respectively, the bottom in very minor amount). Multiple peaks could rise from partial cyclization and/or epimerization. Further work is in progress in order to establish their origin.